Exploring the Effect of Minimum Magnitude on California Seismic Hazard Models

Molly M. Gallahue * 🔟[1], Leah Salditch 🔟[2], Madeleine C. Lucas[3], James Neely 🔟[1,4], Seth Stein 🔟[1,4], Norman Abrahamson[5], Susan E. Hough 🔟[6]

[1]Department of Earth and Planetary Sciences, Northwestern University, Evanston, Illinois U.S.A., [2]U.S. Geological Survey, Geologic Hazards Science Center, Golden, Colorado, U.S.A., [3]Department of Earth and Space Sciences, University of Washington, Seattle, Washington, U.S.A., [4]Institute for Policy Research, Northwestern University, Evanston, Illinois, U.S.A., [5]Department of Civil and Environmental Engineering, University of California, Berkeley, California, U.S.A., [6]U.S. Geological Survey, Earthquake Science Center, Moffett Field, California, U.S.A.

Author contributions: *Conceptualization*: All authors. *Data Curation*: M. Gallahue, L. Salditch, N. Abrahamson, M. Lucas, J. Neely. *Formal Analysis*: M. Gallahue, N. Abrahamson. *Funding Acquisition*: N. Abrahamson, S. Stein, S. Hough. *Investigation*: M. Gallahue, N. Abrahamson, S. Stein, S. Hough. *Methodology*: N. Abrahamson. *Resources*: L. Salditch. *Validation*: M. Gallahue, N. Abrahamson. *Visualization*: M. Gallahue, N. Abrahamson, S. Stein, S. Hough. *Writing – original draft*: M. Gallahue. *Writing – review & editing*: M. Gallahue, S. Stein, S. Hough, L. Salditch.

Abstract A recent topic of interest is the performance of probabilistic seismic hazard maps relative to historically observed shaking. Past studies in several areas have documented an apparent discrepancy between seismic hazard map predictions and historically observed shaking. To explore whether this discrepancy arises because of incompleteness in historical intensity catalogs, we consider maps and historical intensities from California. Current probabilistic seismic hazard maps for California appear to predict stronger short period shaking than historical maxima captured by the California Historical Intensity Mapping Project (CHIMP) dataset between 1857 and 2019. We estimate that CHIMP has a magnitude completeness between M6 and 6.6, whereas California hazard maps assume a minimum magnitude (M_{Min}) of 5. Disaggregating the maps shows that earthquakes smaller than M6 and 6.6 respectively contribute about 25% and 45% of hazard across California. Increasing the hazard map's M_{Min} to 6 and 6.6 respectively reduces the discrepancy between predicted and observed shaking by approximately 10–20% and 30–35%. These reductions are not enough to bring the maps and data in alignment. Thus, M_{Min} inconsistencies contribute to, but are not the sole cause of, the discrepancy between predicted and historically observed shaking. These results may be generally applicable to maps elsewhere, although M_{Min} will vary for different historical datasets.

1 Introduction

Probabilistic seismic hazard analyses (PSHA) forecast the probability, *p,* that during *t* years of observations shaking at a site will exceed the value shown in a model with a τ-year return period, assumed to be described by a negative exponential distribution,

$$p = 1 - \exp\left(\frac{-t}{\tau}\right) \quad (1)$$

(Cornell, 1968; Field, 2008). Assuming the same behavior over time as over space (the ergodic assumption), *p* also gives the fraction of sites where observed shaking is expected to exceed the mapped value for any map return period and duration of observations. As the ratio of the observation time to the return period increases, *p* should also increase because an increasing fraction of the area will have experienced larger earthquakes and thus higher shaking.

This approach, which was introduced by Ward (1995) and used in subsequent analyses (e.g., Albarello and D'Amico, 2008; Fujiwara et al., 2009; Miyazawa and Mori, 2009; Stirling and Gerstenberger, 2010; Nekrasova

et al., 2014; Tasan et al., 2014; Mak and Schorlemmer, 2016a) considers many sites to avoid the constraint that large ground motions at any given site are rare. At any one site, the annual rate of exceedance is $1/\tau$. PSHA models and corresponding maps are based on assumptions about the rates, spatial distribution, and sizes of potential future earthquakes, and ground-motion models (GMMs) that predict the resulting shaking.

Because PSHA maps are critical for earthquake hazard assessment and risk mitigation, assessing how well PSHA models and maps forecast observed shaking is important. The problem is challenging both because of limitations in the available data and because of conceptual issues regarding the assessment of probabilistic forecast performance (Marzocchi and Jordan, 2014; Stein et al., 2015; Wang, 2015; Vanneste et al., 2018; Brooks et al., 2019; Gneiting and Katzfuss, 2014). In particular, shaking data recorded after a PSHA model was made typically span a time period (*t*) that is extremely short compared to the return period of the hazard map (τ) (e.g., 475 or 2475 years). Hence most post-PSHA model shaking datasets do not fully include the moderately large and large earthquakes that often control hazard for an area. One approach to address this problem is retrospective assessment, or hind-

a

CHIMP Data
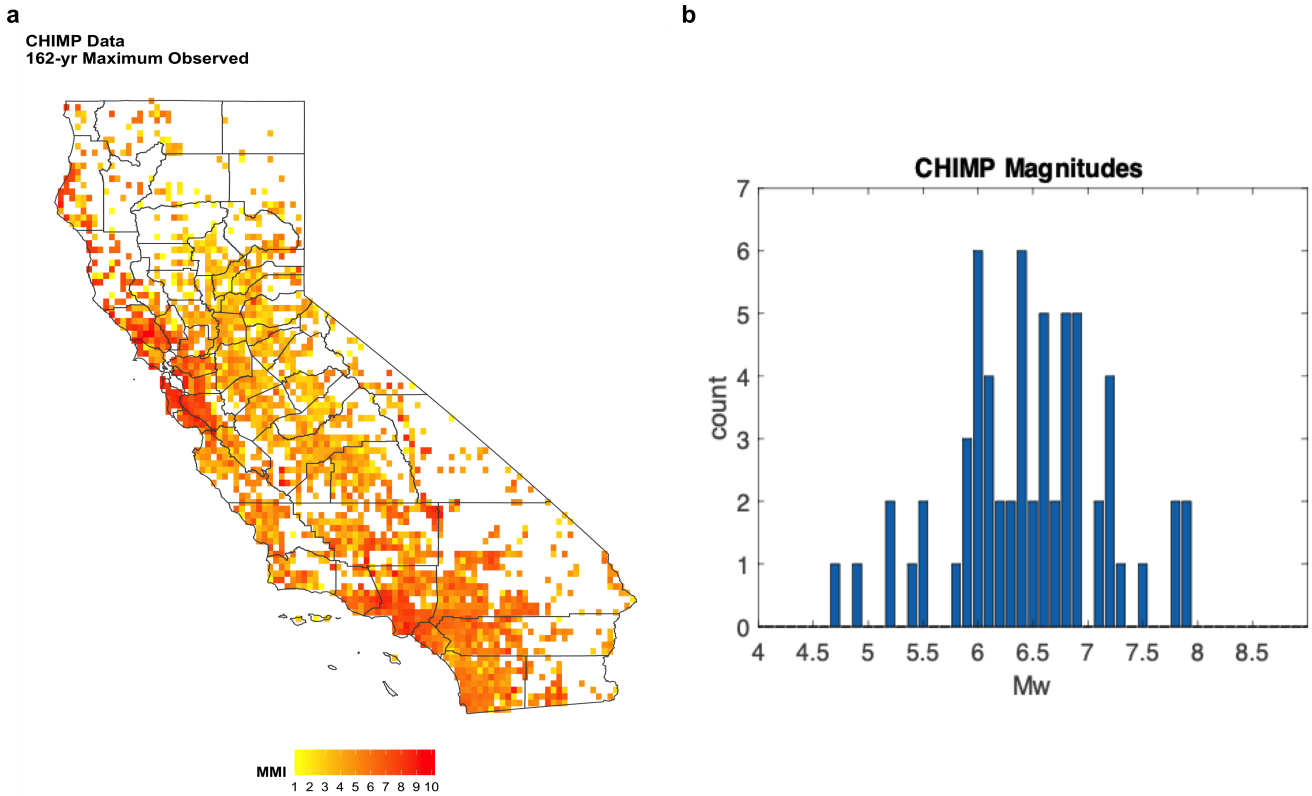**162-yr Maximum Observed**

b

Figure 1 (a) Maximum observed intensity data in 10 × 10 km grid cells from CHIMP (Salditch et al., 2020). (b) Histogram of the magnitudes of the 62 earthquakes included in CHIMP.

casting, which uses compilations of historical shaking data spanning hundreds of years (Stirling and Petersen, 2006; Nekrasova et al., 2014; Mak et al., 2014; Stein et al., 2015; Brooks et al., 2016, 2017, 2018), or in some cases, smaller datasets of instrumental ground-motion data (Stirling and Gerstenberger, 2010; Mak and Schorlemmer, 2016b). The method is similar to that used by meteorologists to improve weather forecasts (Stein et al., 2015).

PSHA maps have recently been compared to historical datasets for several countries. Comparisons of mean hazard maps with catalogs of historical data for Italy, Japan, and California, (USA) have all shown that maps appear to overpredict shaking relative to historical datasets (Stein et al., 2015; Brooks et al., 2016; Salditch et al., 2020). Comparisons for France and Nepal yield similar results (Drouet et al., 2020; Salditch, 2021). Similar studies using ShakeMap data for Australia also demonstrate that the models appear to predict higher shaking than observed in the last 50 years (Allen et al., 2023). Thus, a similar discrepancy arises in different regions for shaking datasets and hazard maps developed by different groups and approaches. Although some studies have found that hazard maps are consistent, or even underpredict, relative to historical datasets (e.g., Rey et al., 2018; Griffin et al., 2019), these studies consider the largest intensities, above which observed data are considered complete. In doing so, these studies exclude regions where observed intensities are significantly less than those predicted, which is where the discrepancy arises. Although a compilation of historically observed shaking may not be complete at the MMI

3 or 4 (say) level, a fair overall comparison of a hazard map should consider grid cells for which one has historically observed maximum intensities. A potential incompleteness bias in historical intensity catalogs is countered by a general reporting bias, whereby severe effects are more likely to be documented than mild or moderate effects (e.g., Hough, 2013). This reporting bias persists with the DYFI system: Mak and Schorlemmer (2016a) showed that, to first order, population density and shaking severity control the likelihood that a DYFI response will be received. Additionally, these studies examine individual sites, which differs from our approach of examining the entire map.

This discrepancy might have several causes, as discussed by Salditch et al. (2020), including that the data may be biased low, the maps may be biased high, or that the misfit may arise purely by chance. One possible cause of the data being biased low is that databases of historically observed shaking may be incomplete for magnitudes that contribute to hazard. To explore this issue, we consider U.S. Geological Survey (USGS) mean probabilistic seismic hazard maps for California and historical observed shaking captured in the California Historical Intensity Mapping Project (CHIMP). The CHIMP catalog is publicly available, and the USGS provides a publicly available tool for disaggregation (https://earthquake.usgs.gov/hazards/interactive/), making this study possible.

CHIMP is the first comprehensive compilation of observed seismic intensities from 62 of the largest earthquakes (moment magnitude **M**4.7-7.9; most with **M**>6) in California over a 162-yr period between 1857 and
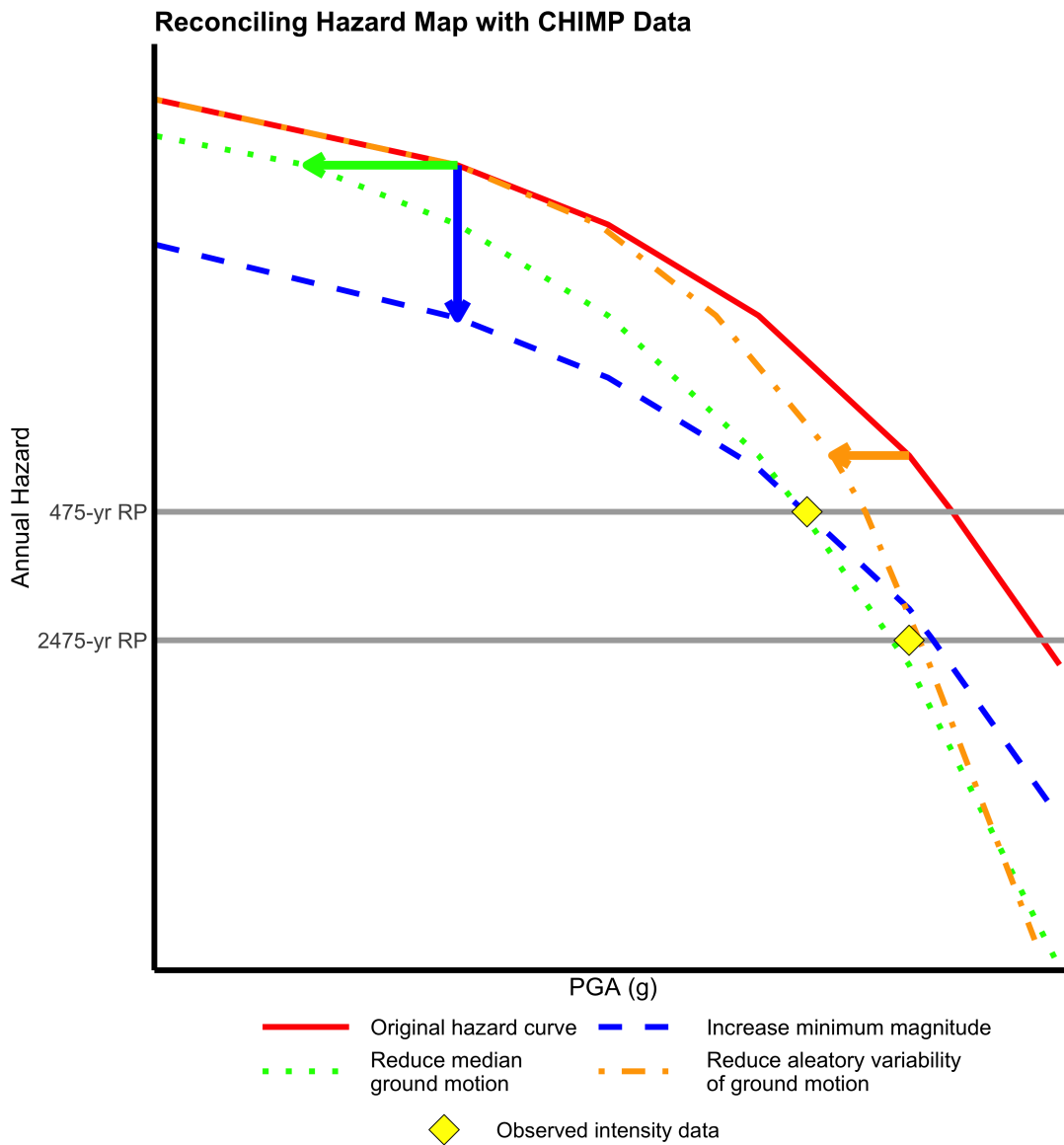
Figure 2 Schematic illustrating three possible adjustments to the hazard model that could make the predicted hazard more consistent with the observed shaking data at a site. Ideally, curves should pass through the yellow diamonds. Curves to the right of the yellow diamonds predict higher shaking than observed in the historical catalog. Arrows show in which direction adjustments change the original hazard curve. 475-yr and 2475-yr return periods (RP) shown.

2019 (Fig. 1b) (Salditch et al., 2020). Intensity values, measured on the Modified Mercalli Intensity (MMI) scale (Wood and Neumann, 1931) were collected from sources including historical newspaper accounts, USGS postcard questionnaires, USGS "Did You Feel It?" (DYFI) program responses, and oral history accounts (Salditch et al., 2020). Assigned intensity values were then aggregated into 10 × 10 km regions, giving the maximum observed intensity in each (Fig. 1a). The nominal uncertainty of each observation is +/- one intensity unit (Salditch et al., 2020), a standard variation between researchers (Hough and Page, 2011; Salditch et al., 2018).

This process is likely to miss shaking due to some smaller historical events for which accounts are limited, and possibly moderate earthquakes outside of California borders that were felt within the state. Salditch et al. (2020) explored CHIMP incompleteness from historical population distributions through interpolated inten-

sities for the three largest earthquakes in California (1857 Fort Tejon, 1872 Owens Valley, and 1906 San Francisco). They showed that results did not change appreciably when more spatially dense intensity values for past large earthquakes were included. Incompleteness of intensity observations at individual points is thus unlikely to account for the discrepancy between maps and observed intensities.

To compare hazard maps with CHIMP, Salditch et al. (2020) converted the 2018 USGS seismic hazard map (Petersen et al., 2020) from peak ground acceleration (PGA) to MMI via Worden et al. (2012) magnitude and distance independent ground-motion intensity conversion equation (GMICE). Here, we use PGA converted to MMI hazard maps for consistency with the Salditch et al. (2020) study.

Equation 1, for $t = 162$ years, predicts that the fraction of sites where observed shaking exceeds the mapped

hazard value should be $p = 0.289$ and 0.063 for 475- and 2475-yr return period maps, respectively. However, the observed fraction of site exceedances, f, calculated by dividing the number of sites where CHIMP data is higher than a 475- and 2475-yr hazard map's prediction by the total number of CHIMP sites, is considerably lower than either p. Salditch et al. (2020) found $f = 0.064$ and 0.006 for the 475- and 2475-yr maps, respectively.

The performance of hazard maps can be quantified using $M0$, the fractional exceedance metric, which is defined as $|f - p|$ (Stein et al., 2015). Ideally this metric would equal 0, and larger $M0$s indicate a greater discrepancy between predictions and observations. We also consider the ratio f/p, which can be thought of as a scale factor for the number of earthquakes contributing to hazard that would be needed to align the map predictions and observations. Map "overprediction" arises when $p > f$.

Salditch et al. (2020) found that p is greater than f for the maps with both a 10% probability of exceedance in 50 years (475-yr return period) and a 2% probability of exceedance in 50 years (2475-yr return period), so both overpredict shaking relative to the CHIMP data. This pervasive discrepancy has important implications and calls for detailed investigation.

The overprediction of the shaking data could be influenced by three different factors, shown schematically via the hazard curve for an individual site (Fig. 2) in a hazard map. If the median ground motions as predicted by GMMs were smaller, the hazard curve would shift horizontally, bringing it in closer alignment with observed data. Alternatively, if the assumed hazard model's $\mathbf{M_{Min}}$, the minimum magnitude of earthquakes included in the hazard calculation, were higher, the hazard curve would shift vertically because hazard is reduced due to fewer contributing earthquakes, again bringing it into closer alignment with observed intensities. Lastly, if the aleatory, or random, variability of the ground-motion models were smaller, the hazard distribution would be narrowed, thus steepening the hazard curve. Other effects can also change the hazard curves, but we consider these to be the simplest and thus most likely.

There are several assumptions made when building and testing a hazard map. The comparisons by Salditch et al. (2020) used PSHA maps for a reference site condition, leaving open the possibility that the results might be significantly different if site response were included. Gallahue et al. (2022) showed, however, that incorporating site effects (as predicted by NGA-West2) did not appreciably change high-frequency PSHA maps, and thus that site effects are not a cause of the discrepancy. The weak predicted influence of site response on high-frequency ground motion is due to nonlinear damping at strong shaking levels. Furthermore, comparison of hazard maps (traditionally given in PGA) to observed shaking records (in seismic intensity) requires the use of a GMICE. Salditch et al. (2020) converted PGA hazard maps to intensity using the Worden et al. (2012) GMICE. However, a recent study by Gallahue and Abrahamson (2023) found that the methodology used in the Worden GMICE (and others) leads to overestimated intensities at

PGAs relevant to hazard maps. These biased GMICE will contribute to the observed discrepancy. USGS hazard maps also rely on the mean hazard, but median hazard curves may better align with observed data. Other factors, such as nonergodic ground-motion models used to develop the maps or spatial correlation effects, could also contribute to the discrepancy.

In this paper, we consider whether removing the contribution to seismic hazard from smaller earthquakes (increasing $\mathbf{M_{Min}}$ in the hazard maps) to adjust for incompleteness of the CHIMP dataset could contribute to the observed overprediction (Fig. 2). This evaluates the assumption that the minimum magnitudes between maps and data are consistent, or the differences are irrelevant. USGS mean PSHA maps, which we refer to as the reference maps, use a minimum magnitude $\mathbf{M_{Min}}$ of 5 in hazard calculations for California (Petersen et al., 2015, 2020). However, the magnitude for completeness is larger for CHIMP. The PSHA maps were also generated using earthquakes within a 100-km zone beyond state borders.
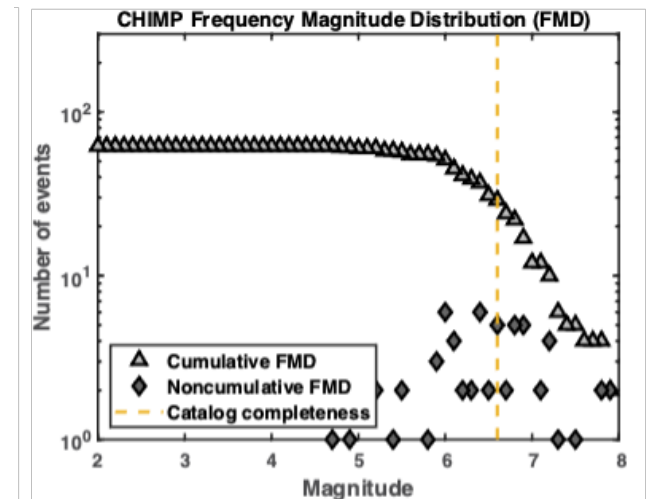

Figure 3 Frequency-magnitude distribution (FMD) of earthquakes in the CHIMP dataset. Estimated catalog completeness level is $\mathbf{M}6.6$ (dashed vertical line).

CHIMP considers all earthquakes greater than $\mathbf{M}6$, and a few smaller ones. Thus, the lowest possible magnitude of completeness of CHIMP is $\mathbf{M}6$. Using the maximum curvature method (Mignan and Woessner, 2012) to calculate the CHIMP catalog's completeness, we find the maximum value of the first derivative of the cumulative frequency-magnitude distribution to be $\mathbf{M}6.6$ (Fig. 3). Hence CHIMP is complete for $\mathbf{M}6.6+$ but includes some smaller events if they were suspected to control the maximum observed intensity at some sites. We further note that evaluations of PSHA maps using CHIMP rely on locations for which historical intensities are available. Catalog completeness will be lowest in sparsely populated regions; i.e., intensity values will tend to be missing in regions where the catalog is least complete. We thus expect that missing moderate earthquakes in sparsely populated regions, including in the 100-km zone outside of California borders, will tend to

not contribute to assessed map performance. We thus assume that the actual magnitude of completeness for CHIMP is somewhere between **M**6 and 6.6 and evaluate M_{Min} at these end members. If M_{Min} = 6, 51 earthquakes remain in the catalog. If M_{Min} = 6.6, 28 earthquakes remain. For completeness, because M_{Min} can be difficult to resolve, we also tested M_{Min} = 6.8. These results did not differ from those for M_{Min} = 6.6.

2 Disaggregation of hazard for M>6 and M>6.6

We explore the effect of the minimum magnitude using disaggregation, which gives the fractional contribution to the total predicted hazard for a range of magnitudes and distances. Disaggregation results allow the contribution to hazard from varying magnitudes, distances, and epsilons (number of standard deviations above the median ground motion) to be understood (Bazzurro and Cornell, 1999). Effectively, this breaks the hazard map into the input scenarios, and determines the effect of each. From a mathematical basis, the fraction of the hazard from all events with **M**≥5 that results from events with $M \geq M_{Min}$ is:

$$
\begin{aligned}
&\text{Haz}_{PGA}\left(z | M \geq M_{Min}\right) \\
&= \text{Haz}_{PGA}(z | M \geq 5)\,\text{Disagg}\left(M \geq M_{Min} | PGA > z\right)
\end{aligned}
$$

(2)

where z is predicted ground motion (in PGA), Disagg(M≥M_{Min}|PGA>z) is the hazard disaggregation assuming a magnitude greater than M_{Min} and PGA greater than z, and Haz$_{PGA}$(z|M≥ M_{Min}) is the hazard in terms of PGA of z assuming M≥5. Figure 4, a disaggregation result for one site in California, demonstrates how the removal of magnitudes less than a certain value would reduce the total hazard (rate of exceedance) due to using a larger M_{Min} in the hazard calculations. Excluding **M**5–6 reduces the hazard less than excluding **M**5–6.6 because fewer events are excluded.

We produce approximate hazard maps for M_{Min}= 6 and M_{Min} =6.6 using disaggregation results for a grid of locations in California using the USGS Unified Hazard Tool for return periods of 475 years and 2475 years, corresponding to the two maps examined by Salditch et al. (2020). Because disaggregation reports for the 2018 mean hazard model and map were not available in the USGS tool at the time of this analysis, we used the 2014 mean values. Petersen et al. (2020) found that the mean hazard in California stayed consistent between the 2014 model and the 2018 update, so results for 2014 are expected to be comparable to the more recent 2018 models. Additionally, the earthquake catalog and source models for California in both the 2014 and 2018 maps are from the Uniform California Earthquake Rupture Forecast model (UCERF3). Thus, no earthquakes were added or removed from 2014 to 2018, and no substantial changes occurred in California hazard during this time (Petersen et al., 2020, 2021).

We compute the fraction of the hazard due to larger events at each grid site. The disaggregation shows that the average contribution to hazard of **M**≥6 events across
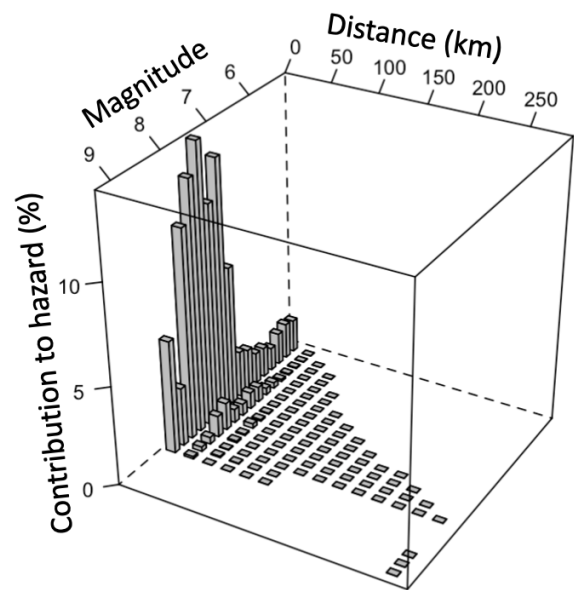
Contribution from all sources

Figure 4 Disaggregation results for a site outside San Francisco. Most of the hazard comes from events with magnitudes greater than ~7 and distances < ~50 km. If the contribution from smaller magnitude events were removed, the total hazard would decrease.

all sites is ~75% for the 475-yr return period map and ~76% for the 2475-yr return period map. Similarly, the average contribution to hazard of **M**≥6.6 events across all sites is smaller, ~53% for the 475-yr return period map and ~55% for the 2475-yr return period map. At each grid site, we reduce the hazard to that excluding **M**<6 or 6.6 by scaling the hazard curve (Rukstales and Petersen, 2019) by the calculated percent contribution to hazard. Via log-interpolation on the hazard curves, we derive the corresponding decrease in PGA. To scale the hazard maps for the hazard from only **M**≥6 or 6.6 events, we adjust the reference hazard map PGA by the PGA percent decrease at the closest grid location (Fig. 5).

The percent decreases in PGA from the reference map to the maps scaled to exclude **M**<6 or 6.6 are shown in Fig. 6. Excluding the contribution from smaller earthquakes reduces the hazard and thus the corresponding predicted PGA. For 475-yr and 2475-yr return periods, the reduction in the PGA from excluding **M**<6 events ranges from less than 10% near the major coastal faults to up to 35% in the Sierra foothills. Excluding the contribution from **M**<6.6 events reduces PGA by up to 65% in the Sierra foothills.

These results show that **M**<6 events contribute substantially to predicted hazard (~25%), and **M**<6.6 events contribute even more (~45%). This result is consistent with the result of Minson et al. (2009), that due to the aleatory variability of ground motions and the Gutenberg-Richter distribution of magnitudes (Gutenberg and Richter, 1944), strong short-period shaking is more likely to be generated by more frequent moderate-
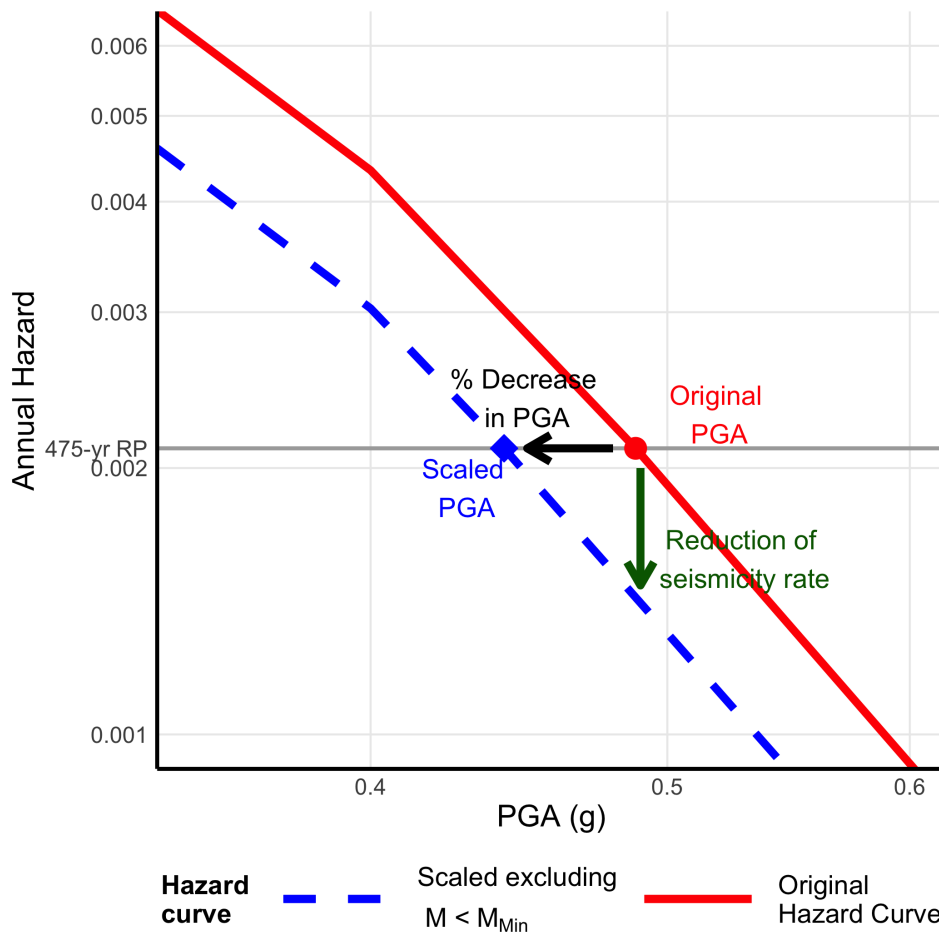
Figure 5 Schematic of decrease in hazard (PGA) with a 475-yr return period (RP) resulting from the hazard curve reduction from increasing M_{Min}.

magnitude earthquakes than by rarer large-magnitude earthquakes. Fig. 6 shows that **M**<6 and 6.6 events strongly contribute to hazard in some areas. Conceptually, we expect that hazard near major, active faults will be controlled by large earthquakes on those faults. However, away from major faults, hazard will be more controlled by distributed background seismicity, so relatively frequent moderate-magnitude events will contribute more to overall hazard. Figure 7 shows the decrease of the scaled map relative to the reference for all cases, highlighting the MMI reduction from removing the contribution to hazard from smaller earthquakes.

It is worth noting that our described approach produces approximate maps for the given M_{Min}. The approach used here assumes that the relative contributions of the sources to the target exceedance rate does not appreciably change between the original and scaled maps. To completely alter M_{Min} would require changing the minimum magnitude during the calculation of hazard curves that underlie the maps. In doing so, the median and variance of ground motion and the activity rate, all of which influence predictions, would change. As described in Figure 2, changes to the median ground motion shift the hazard curve horizontally, whereas changes to variance would affect the slope of the hazard curve. Changes to the activity rate would shift the hazard curve vertically (Figure 2). However, developing hazard maps involves extensive computations, so in practice it is often not justifiable to compute hazard models for multiple M_{Min}, as evaluated in this paper.

3 Comparison of M≥6 Or 6.6 Maps with Historical Intensity Data

Using the metrics from Salditch et al. (2020), we assessed how the hazard maps scaled for **M**≥6 or 6.6 perform relative to CHIMP historical intensity data. For this comparison, we edited the CHIMP dataset to include data from only **M**≥6 or 6.6 earthquakes, which reduced the number of sites from that shown in Fig. 1a. We only assess hazard map values at the locations in which CHIMP data existed in the **M**≥6 or 6.6 sets; thus, the total number of sites evaluated varies between each considered M_{Min}. The **M**≥6 or 6.6 scaled PGA map values were converted to MMI via Worden et al. (2012)'s magnitude- and distance- independent equation that does not include aleatory variability.

Comparison of exceedance rates (Fig. 8, 9) shows that the scaled hazard maps excluding **M**<6 or 6.6 perform better than the reference maps because the lower predicted hazard yields more observed exceedances (Table 1, 2). However, the fractions, *f*, of sites where the largest shaking exceeds that predicted from either map remain much lower than the predicted fraction, *p*.

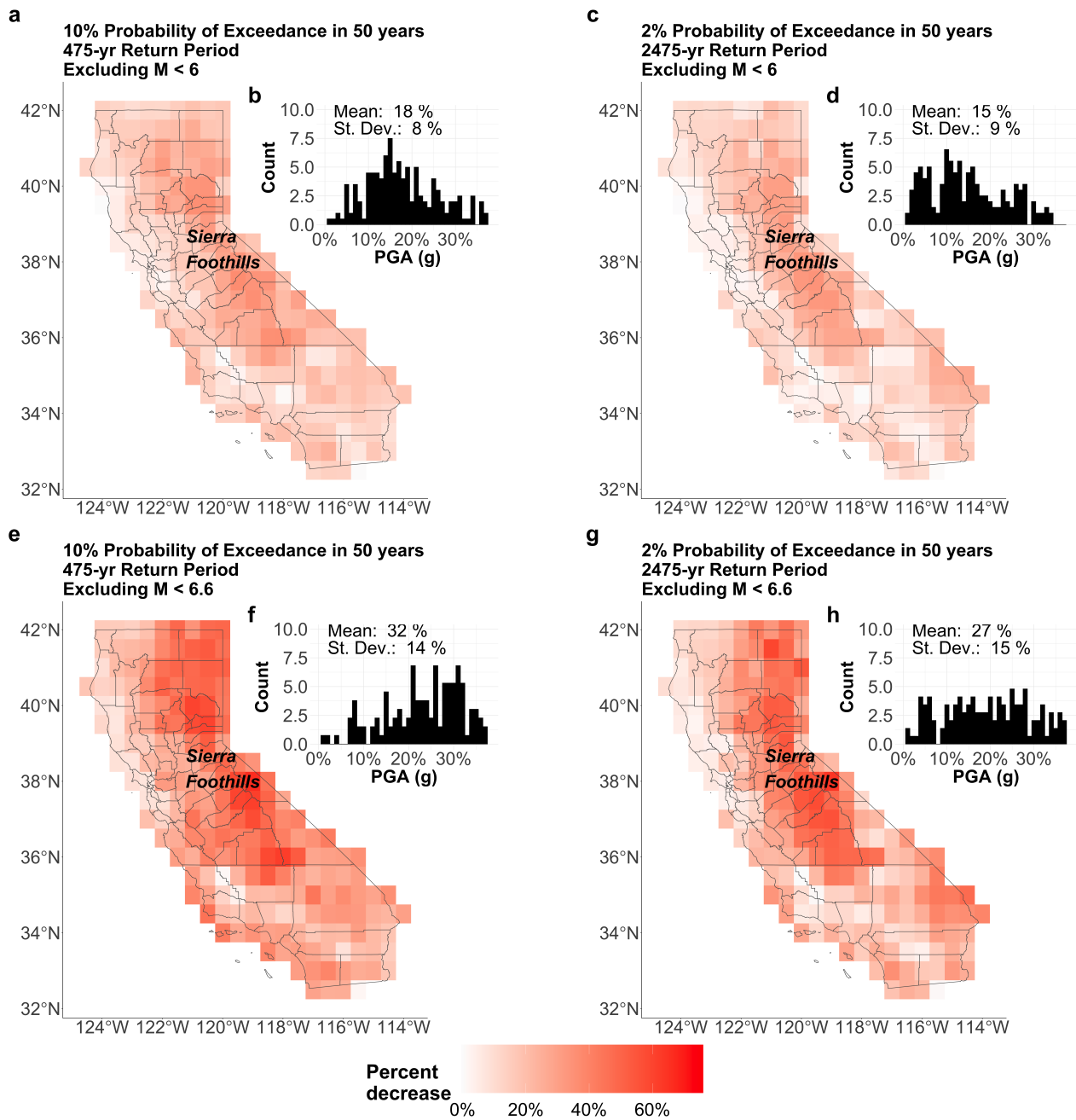## Percent Decrease in PGA Excluding Hazard from M < M_Min

Figure 6 Maps showing percent decreases in PGA. For 10% probability of exceedance in 50 years (475-yr return period) map (a) excludes **M**<6 earthquakes and (e) excludes **M**<6.6 earthquakes. For the 2% probability of exceedance in 50 years (2475-yr return period) map (c) excludes **M**<6 earthquakes and (g) excludes **M**<6.6 earthquakes. Distributions of the percent decreases for (b), (f) the 475-yr return period and (d), (h) the 2475-yr return period maps. PGA decreases by an average of 18% for the 475-yr return period map and 15% for the 2475-yr return period map when excluding **M**<6. PGA decreases by an average of 32% for the 475-yr return period map and 27% for the 2475-yr return period map when excluding **M**<6.6. The most substantial PGA decrease is in the Sierra foothills.

For the 475-yr map with $p = 0.289$ when excluding M<6, f increases from 0.071 (155 exceedances) to 0.107 (234 exceedances) for the scaled map. This represents an increase of ~51%, or 79 sites. Similarly, for the 2475-yr map with $p = 0.063$, f increases from 0.01 (21 exceedances) to 0.016 (35 exceedances), an increase of ~67%, or 14 sites. As a result, $M0$ values for the scaled maps were 0.182 and 0.047 for the 10% probability and the 2% probability of exceedance in 50 years maps, lower (showing better fit) than the reference maps' val-

ues of 0.218 and 0.053. The f/p ratios increased from 0.246 and 0.159 for the 475-yr and 2475-yr reference maps to 0.370 and 0.254 for the scaled maps, respectively, indicating improvement in map performance.

Similarly, for the 475-yr map with $p = 0.289$ when excluding **M**<6.6, f increases from 0.063 (128 exceedances) to 0.139 (283 exceedances) for the scaled map. This represents an increase of ~121%, or 155 sites. For the 2475-yr map with $p = 0.063$, f increases from 0.009 (19 exceedances) to 0.025 (50 exceedances), an increase of
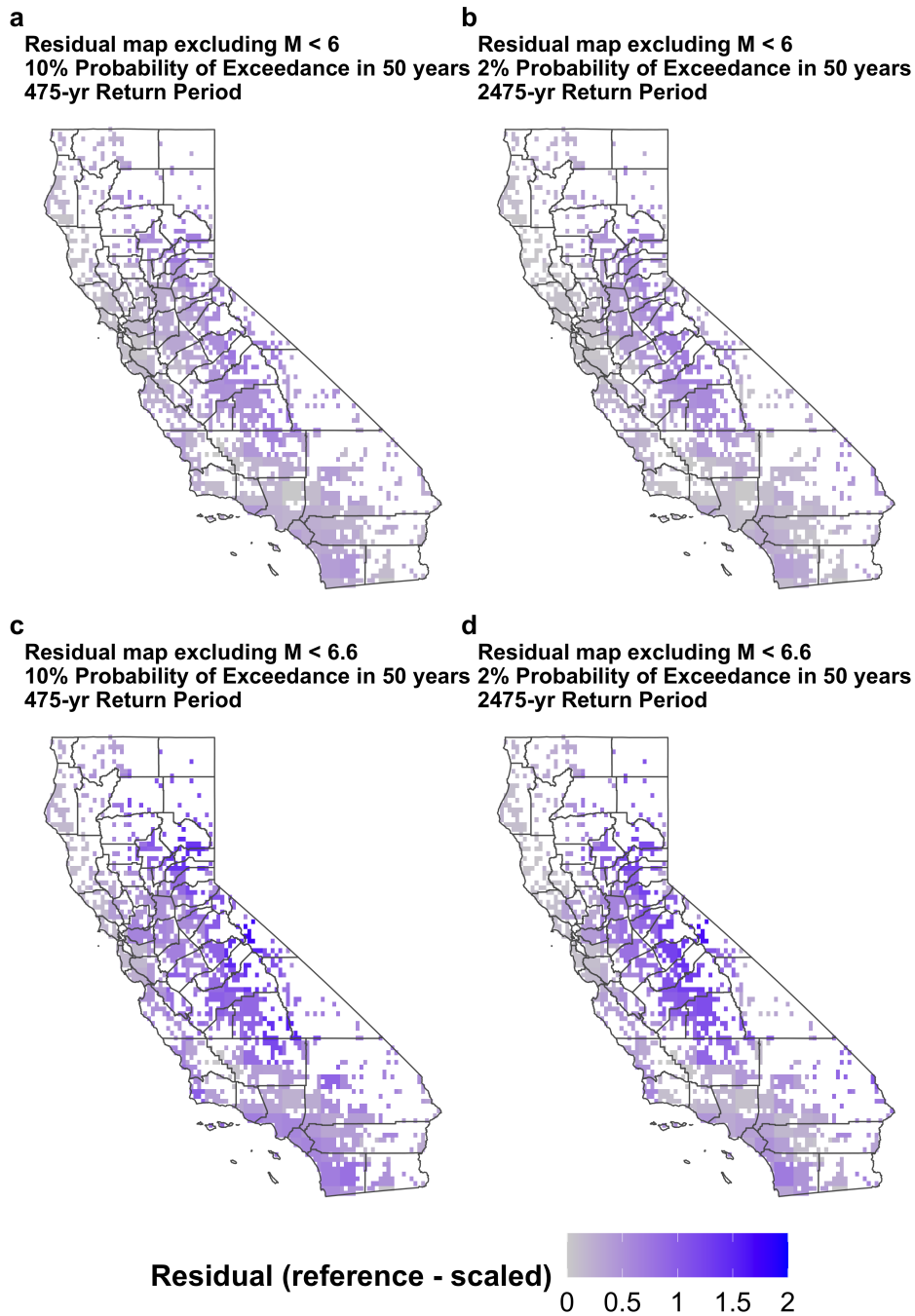
a
Residual map excluding M < 6
10% Probability of Exceedance in 50 years
475-yr Return Period

b
Residual map excluding M < 6
2% Probability of Exceedance in 50 years
2475-yr Return Period

c
Residual map excluding M < 6.6
10% Probability of Exceedance in 50 years
475-yr Return Period

d
Residual map excluding M < 6.6
2% Probability of Exceedance in 50 years
2475-yr Return Period

Residual (reference - scaled)

0   0.5   1   1.5   2

Figure 7   Intensity residuals (reference map intensity minus scaled map intensity) from this study. Scaled map excluding hazard **M** < 6 decreases the 475-yr return period (a) and 2475-yr return period (b) mapped values moderately relative to the reference mapped values. Scaled map excluding hazard **M** < 6.6 decreases the 475-yr return period (c) and 2475-yr return period (d) mapped values even more relative to the reference mapped values.

~178%, or 31 sites. M0 values for the scaled maps, 0.15 and 0.038 for the 10% probability and the 2% probability of exceedance in 50 years map, were lower (showing better fit) than those for the reference maps, 0.226 and 0.054. Similarly, *f/p* ratios increased from 0.218 and 0.143 for the 475-yr and 2475-yr reference maps to 0.481 and 0.397 for the scaled maps, respectively, indicating improvement in map performance.

Despite the improvements, the scaled maps still over-predict shaking relative to CHIMP data. We construct 5–95% intervals for the predicted fraction of exceedance to explore if this could be due to chance. We assume the uncertainty in the number of observed ex-

ceedances in 162 years follows a Poisson probability model, such that

$$P(m \leq N) = e^{\frac{-t}{\tau}} \sum_{m=1}^{N} \frac{\frac{t}{\tau}^{m}}{m!} \qquad (3)$$

where t is the length of the interval (162 years), τ is approximately the average number of years between each exceedance (equal to 162/N), N is the predicted number of exceedances, equal to 588 for the 475-yr maps and 128 for the 2475-yr return-period maps. We compute 5% and 95% uncertainty ranges for N by finding $N_{0.05}$ and $N_{0.95}$ such that $P(m \leq N_{0.05})=0.05$ and $P(m \leq N_{0.95})=0.95$.

	Reference map: 475-yr return period	Scaled M6 map: 475-yr return period	Reference map: 2475-yr return period	Scaled M6 map: 2475-yr return period
p	0.289	0.289	0.063	0.063
5-95% range on p	[0.270, 0.308]	[0.270, 0.308]	[0.054, 0.072]	[0.054, 0.072]
f	0.071	0.107	0.01	0.016
Predicted number of exceedances	632	632	138	138
Observed number of exceedances	155	234	21	35
M0 (ideally 0)	0.218	0.182	0.053	0.047
f/p (ideally 1)	0.246	0.370	0.159	0.254

Table 1 Statistics of comparison of CHIMP data for $\mathbf{M}\geq 6$ to both reference and scaled ($\mathbf{M}\geq 6$) maps.

	Reference map: 475-yr return period	Scaled M6.6 map: 475-yr return period	Reference map: 2475-yr return period	Scaled M6.6 map: 2475-yr return period
p	0.289	0.289	0.063	0.063
5-95% range on p	[0.270, 0.309]	[0.270, 0.309]	[0.054, 0.072]	[0.054, 0.072]
f	0.063	0.139	0.009	0.025
Predicted number of exceedances	588	588	128	128
Observed number of exceedances	128	283	19	50
M0 (ideally 0)	0.226	0.15	0.054	0.038
f/p (ideally 1)	0.218	0.481	0.143	0.397

Table 2 Statistics of comparison of CHIMP data for $\mathbf{M}\geq 6.6$ to both reference and scaled ($\mathbf{M}\geq 6.6$) maps.

Then, $N_{0.05}/2186$ and $N_{0.95}/2186$ are the lower and upper ranges of p, respectively, for the 2186 total sites for $\mathbf{M}\geq 6$ comparisons (Fig. 8). Likewise, $N_{0.05}/2033$ and $N_{0.95}/2033$ are the lower and upper ranges of p, respectively, for the 2033 total sites for $\mathbf{M}\geq 6.6$ (Fig. 9).

For $\mathbf{M}\geq 6$ maps with 2186 sites and predicted exceedances at 632 and 155 sites, the 5–95% uncertainty ranges on p are [0.270, 0.308] and [0.054, 0.072] for the 475-yr and 2475-yr return period maps, respectively. For $\mathbf{M}\geq 6.6$ maps with 2033 sites and predicted exceedances at 588 and 128 sites, the 5–95% uncertainty ranges on p are [0.270, 0.309] and [0.054, 0.072] for the 475-yr and 2475-yr return period maps. In all cases, the observed fraction of exceedances, f, for both the reference and scaled maps fall outside the uncertainty ranges of the predicted fraction, p. Thus, we conclude that the observed remaining discrepancy is not due to chance.

4 Discussion

Our results show that the apparent overpredictions of the USGS mean hazard maps relative to the CHIMP dataset cannot be fully explained by increasing $\mathbf{M_{Min}}$ to 6 or 6.6 to account for incompleteness of CHIMP. Although removing the contribution to hazard due to $\mathbf{M}<6$ or 6.6 events reduces the overprediction, it does not reduce it enough to fall within 5–95% confidence intervals on the predicted fraction of exceedance for either return period.

It remains possible that the discrepancies between predicted and historically observed ground motions are due to long-term incompleteness of the CHIMP catalog due to the short length of the historical record. If, however, the inconsistencies between the historical magnitude of completeness and hazard map $\mathbf{M_{Min}}$ for California were the only cause of the discrepancy, the number of earthquakes contributing to hazard would need to be reduced by approximately a factor of 4 to align the reference map's (including hazard from $\mathbf{M}<6.6$) observed fraction of exceedance with the 5–95% confidence intervals of the predicted for the 475-yr map (5% interval of f/p for reference maps). Similarly, approximately a factor of 6 reduction would be needed for the 2475-yr map. For the scaled maps excluding the hazard from $\mathbf{M}<6.6$ events, reductions of about a factor of 2 for the 475-yr return period map and for the 2475-yr return period map would be needed to bring the metrics in alignment. Even greater factors would be needed to bring the scaled maps excluding hazards from $\mathbf{M}<6$ maps in alignment. Although Page and Felzer (2015) show that seismicity rate tends to be underestimated from short catalogs, their results indicate it is unlikely that the rate has been underestimated by such large factors. Recent results compiled by the U.S. Geological Survey "Did You Feel It?" system reveal that shaking from $\mathbf{M}6$ earthquakes is widely felt. For example, the 8 July 2021 $\mathbf{M}6$ Antelope Valley earthquake, which struck a remote area near the Nevada border, was widely felt throughout north-central California (see Data & Code Availability and Reproducibility).
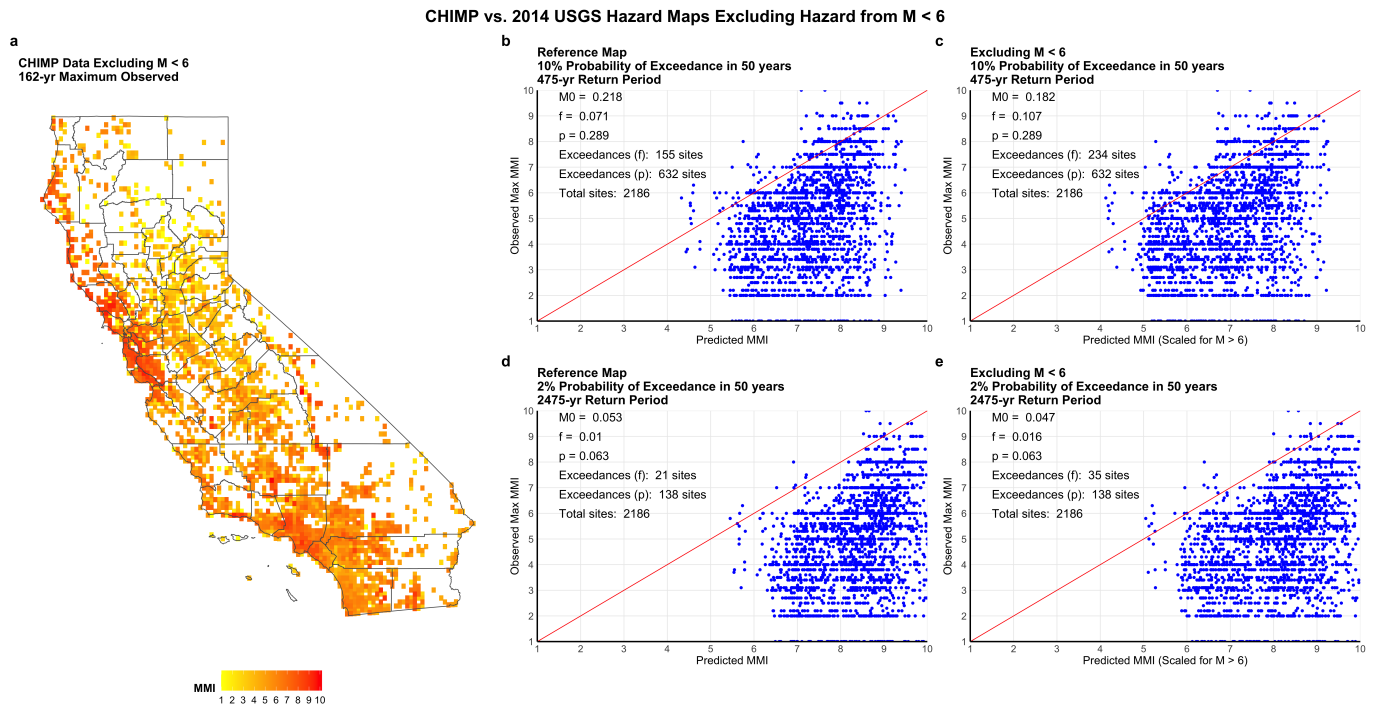
Figure 8 (a) CHIMP data excluding **M**<6 events. (b–e) Predicted versus observed shaking for the 2014 reference USGS hazard map with 10% probability of exceedance in 50 years (b) and scaled hazard map excluding **M**<6 (c). (d) and (e) show corresponding results for the map with 2% probability of exceedance in 50. 1:1 line shown in red.
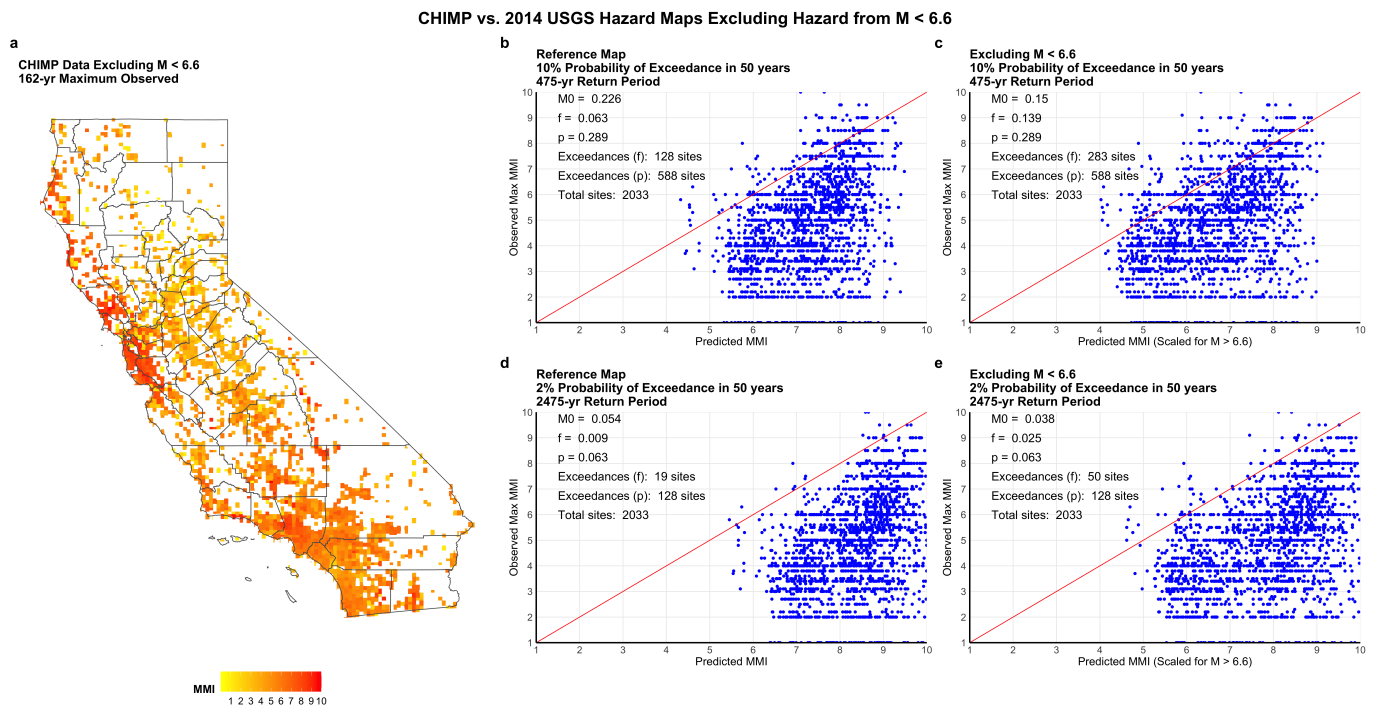

Figure 9 (a) CHIMP data excluding **M**<6.6 events. b–e) Predicted vs. observed shaking for the 2014 reference USGS hazard map with 10% probability of exceedance in 50 years (b) and scaled hazard map excluding **M**<6.6 (c). (d) and (e) show corresponding results for the map with 2% probability of exceedance in 50. 1:1 line shown in red.

5 Conclusion

We have considered one possible factor that might explain the discrepancy between predicted shaking in California and the much lower historically observed shaking revealed by the CHIMP dataset (Salditch et al., 2020). Our analysis shows that **M**5–6 earthquakes contribute

approximately 25% of the predicted hazard in California and **M**5–6.6 contribute about 45%, consistent with implications of Minson et al. (2009). Hence, when the hazard maps are scaled to include only the effects of **M** >6 or 6.6 events, the observed overpredictions lessen. However, this decrease is not large enough that the observed fractional exceedance falls within the 5–95%

confidence interval of that predicted. As such, the discrepancy between predicted and historically observed shaking in California is not due to incompleteness of the CHIMP dataset. These inconsistencies in minimum magnitude contribute to the discrepancy but are less than 30–35% of the discrepancy. Analogous results are expected to arise for discrepancies between maps and historical datasets in other areas when their corresponding M_{Min}s are considered. Future work is planned to focus on other possible effects contributing to the overprediction, including improved conversion equations between PGA and MMI (Gallahue and Abrahamson, 2023), which could lead to improved seismic hazard maps for California and worldwide.

Acknowledgements

Data and code availability

The CHIMP dataset is available as an electronic supplement to Salditch et al. (2020) and at https://github.com/salditch/CHIMP. The USGS hazard maps were downloaded from Rukstales and Petersen (2019) at https://doi.org/10.5066/P9WT5OVB (last accessed October 2021). The 2014 hazard curves and maps were downloaded from Shumway (2019) at https://doi.org/10.5066/P9P77LGZ (last accessed 28 October 2021 and 15 November 2021, respectively). Disaggregation files were downloaded from the Unified Hazard Tool, accessed at https://earthquake.usgs.gov/hazards/interactive/ from the Dynamic: Conterminous U.S. 2014 (v4.1.4) edition (last accessed 3 November 2021). 2014 USGS hazard maps were downloaded from https://www.sciencebase.gov/catalog/item/5db9be62e4b06957974eb5ca, last accessed 15 November 2021. MATLAB scripts for the catalog completeness calculations are from "Basic Statistical Seismology Tutorial" by Brendan Sullivan, modified by Zhigang Peng, were downloaded from http://geophysics.eas.gatech.edu/people/bsullivan/tutorial/StatisticalSeismology.htm#part1_3 (last accessed 21 July 2022). Did You Feel It? results for the 2021 Antelope Valley earthquake can be accessed at https://earthquake.usgs.gov/earthquakes/eventpage/nc73666231/dyfi/intensity (last accessed 16 August 2022).

Competing interests

The authors declare no competing interests.

References

Albarello, D. and D'Amico, V. Testing probabilistic seismic hazard estimates by comparison with observations: An example in Italy. *Geophysical Journal International*, 175(3), 2008. doi: 10.1111/j.1365-246X.2008.03928.x.

Allen, T., Ghasemi, H., and Griffin, J. Exploring Australian hazard map exceedance using an Atlas of historical ShakeMaps. Earthquake Spectra, 2023.

Bazzurro, P. and Cornell, A. Disaggregation of seismic hazard. *Bulletin of the Seismological Society of America*, 89(2), 1999. doi: 10.1785/BSSA0890020501.

Brooks, E., Stein, S., and Spencer, B. Comparing the performance of Japan's earthquake hazard maps to uniform and randomized maps. *Seismological Research Letters*, 87(1), 2016. doi: 10.1785/0220150100.

Brooks, E., Stein, S., and Spencer, B. Investigating the effects of smoothing on the performance of earthquake hazard maps. *International Journal of Earthquake and Impact Engineering*, 2(2), 2017. doi: 10.1504/IJEIE.2017.10010009.

Brooks, E., Stein, S., Spencer, B., Salditch, L., Petersen, M., and D.E, M. Assessing earthquake hazard map performance for natural and induced seismicity in the central and eastern United States. *Seismological Research Letters*, 89(1), 2018. doi: 10.1785/0220170124.

Brooks, E., Neely, J., Stein, S., Spencer, B., and Salditch, L. Assessments of the performance of the 2017 one-year seismic hazard forecast for the Central and Eastern United States via simulated earthquake shaking data. *Seismological Research Letters*, 90(3), 2019. doi: 10.1785/0220190007.

Cornell, A. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58(5), 1968. doi: 10.1785/B-SSA0580051583.

Drouet, S., Ameri, G., Le Dortz, K., Secanell, R., and Senfaute, G. A probabilistic seismic hazard map for metropolitan France. *Bulletin of Earthquake Engineering*, 18(3), 2020. doi: 10.1007/s10518-020-00790-7.

Field, E. Probabilistic seismic hazard analysis: A primer, 2008. https://opensha.org/resources/PSHA_Primer_v2_0.pdf.

Fujiwara, H., Morikawa, N., Ishikawa, Y., Okumura, T., and J, M. Statistical comparison of national probabilistic seismic hazard maps and frequency of recorded JMA seismic intensities from the K-NET strong-motion observation network in Japan during 1997-2006. *Seismological Research Letters*, 80(3), 2009. doi: 10.1785/gssrl.80.3.458.

Gallahue, M. and Abrahamson, N. New Methodology for Unbiased Ground-Motion Intensity Conversion Equations (GMICE. *Bulletin of the Seismological Society of America*, 2023. doi: 10.1785/0120220224.

Gallahue, M., Salditch, L., Lucas, M., Neely, J., Stein, S., Abrahamson, N., Williams, T., and Hough, S. A study on the effect of site response on California seismic hazard map assessment. *Frontiers in Earth Science*, 10, 2022. doi: 10.3389/feart.2022.931340.

Gneiting, T. and Katzfuss, M. Probabilistic forecasting. *Annual Review of Statistics and Its Application*, 1, 2014. doi: 10.1146/annurev-statistics-062713-085831.

Griffin, J., Nguyen, N., Cummins, P., and Cipta, A. Historical earthquakes of the eastern Sunda arc: Source mechanisms and intensity-based testing of Indonesia's National Seismic Hazard Assessment. Bulletin of the Seismological Society of America, 109(1, 2019.

Gutenberg, B. and Richter, C. Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4), 1944.

doi: 10.1785/BSSA0340040185.

Hough, S. Spatial variability of "Did You Feel It?" intensity data: Insights into sampling biases in historical earthquake intensity distributions. *Bulletin of the Seismological Society of America*, 103(5), 2013. doi: 10.1785/0120120285.

Hough, S. and Page, M. Toward a consistent model for strain accrual and release for the New Madrid Seismic Zone, central United States. *Journal of Geophysical Research*, 116, B03311, 2011. doi: 10.1029/2010JB007783.

Mak, S. and Schorlemmer, D. What makes people respond to "Did You Feel It?". *Seismological Research Letters*, 87(1), 2016a. doi: DOI:10.1785.0220150056.

Mak, S. and Schorlemmer, D. A comparison between the forecast by the U.S. National Seismic Hazard Maps with recent ground-motion records. *Bulletin of the Seismological Society of America*, 106(4), 2016b. doi: 10.1785/0120150323.

Mak, S., Clements, R., and Schorlemmer, D. The statistical power of testing probabilistic seismic-hazard assessments. *Seismological Research Letters*, 85(4), 2014. doi: 10.1785/0220140012.

Marzocchi, W. and Jordan, T. Testing for ontological errors in probabilistic forecasting models of natural systems. *Proceedings of the National Academy of Science of the United States of America*, 111(33), 2014. doi: 10.1073/pnas.1410183111.

Mignan, A. and Woessner, J. Estimating the magnitude of completeness for earthquake catalogs, Community Online Resource for Statistical Seismicity Analysis, 45, 2012.

Minson, S., Baltay, A., Cochran, E., McBride, S., and Milner, K. Shaking is almost always a surprise: The earthquakes that produce significant ground motion. *Seismological Research Letters*, 92 (1), 2009. doi: 10.1785/0220200165.

Miyazawa, M. and Mori, J. Test of seismic hazard map from 500 years of recorded intensity data in Japan. *Bulletin of the Seismological Society of America*, 99(6), 2009. doi: 10.1785/0120080262.

Nekrasova, A., Kossobokov, V., Peresanand, A., and Magrin, A. Comparison of NDSHA, PSHA seismic hazard maps and real seismicity for the Italian territory. *Natural Hazards*, 70(1), 2014. doi: 10.1007/s11069-013-0832-6.

Page, M. and Felzer, K. Southern San Andreas fault seismicity is consistent with the Gutenberg-Richter magnitude-frequency distribution. *Bulletin of the Seismological Society of America*, 105 (4), 2015. doi: 10.1785/0120140340.

Petersen, M., Moschetti, M., Powers, P., Mueller, C., Haller, K., Frankel, A., Zeng, Y., Rezaeian, S., Harmsen, S., Boyd, O., and Field, N. The 2014 United States National Seismic Hazard Model. *Earthquake Spectra, 31(1*, 2015. doi: 10.1193/120814EQS210M.

Petersen, M., Shumway, A., Powers, P., Mueller, C., Moschetti, M., Frankel, A., Rezaeian, S., McNamara, D., Luco, N., Boyd, O., Rukstales, K., Jaiswal, K., Thompson, E., Hoover, S., Clayton, B., Field, E., and Zeng, Y. The 2018 update of the US National Seismic Hazard Model: Overview of model and implications. *Earthquake Spectra*, 36(1), 2020. doi: 10.1177/8755293019878199.

Petersen, M., Shumway, A., Powers, P., Mueller, C., Moschetti, M., Frankel, A., Rezaeian, S., McNamara, D., Luco, N., Boyd, O., Rukstales, K., Jaiswal, K., Thompson, E., Hoover, S., Clayton, B., Field, E., and Zeng, Y. The 2018 update of the US National Seismic Hazard Model: Where, why, and how much probabilistic ground motion maps changed. *Earthquake Spectra*, 37(2), 2021. doi: 10.1177/8755293020988016.

Rey, J., Beauval, C., and Douglas, J. Do French macroseismic intensity observations agree with expectations from the European Seismic Hazard Model 2013? *Journal of Seismology, 22(3*, 2018. doi: 10.1007/210950-017-9724-7.

Rukstales, K. and Petersen, M. Data release for 2018 update of the U.S. National Seismic Hazard Model, U.S. *Geological Survey data release*, 2019. doi: 10.5066/P9WT5OVB.

Salditch, L. *Investigating Earthquake Recurrence and Hazard Models [Doctorate dissertation*. Northwestern University, 2021. doi: 10.21985/n2-my3w-v278.

Salditch, L., Hough, S., Stein, S., Spencer, B., Brooks, E., Neely, J., and Lucas, M. The 1952 Kern County, California earthquake: A case study of issues in the analysis of historical intensity data for estimation of source parameters. *Physics of the Earth and Planetary Interiors. In*, 283, 2018. doi: 10.1016/j.pepi.2018.08.007.

Salditch, L., Gallahue, M., Lucas, M., Neely, J., Hough, S., and Stein, S. California Historical Intensity Mapping Project (CHIMP): A consistently reinterpreted dataset of seismic intensities for the past 162 yr and implications for seismic hazard maps. *Seismological Research Letters*, 91(5), 2020. doi: 10.1785/0220200065.

Shumway, A. Data release for the 2014 National Seismic Hazard Model for the conterminous U.S., U.S. *Geological Survey data release*, 2019. doi: 10.5066/P9P77LGZ.

Stein, S., Spencer, B., and Brooks, E. Metrics for assessing earthquake hazard map performance. *Bulletin of the Seismological Society of America*, 105(4), 2015. doi: 10.1785/0120140164.

Stirling, M. and Gerstenberger, M. Ground motion-based testing of seismic hazard models in New Zealand. *Bulletin of the Seismological Society of America*, 100(4), 2010. doi: 10.1785/0120090336.

Stirling, M. and Petersen, M. Comparison of the historical record of earthquake hazard with seismic-hazard models for New Zealand and the continental United States. *Bulletin of the Seismological Society of America*, 96(6), 2006. doi: 10.1785/0120050176.

Tasan, H., Beauval, C., Helmstetter, A., Sandikkaya, A., and Guéguen, P. Testing probabilistic seismic hazard estimates against accelerometric data in two countries: France and Turkey. *Geophysical Journal International*, 198(3), 2014. doi: 10.1093/gji/ggu191.

Vanneste, K., Stein, S., Camelbeeck, T., and Vleminckx, B. Insights into earthquake hazard maps performance from shaking history simulations. *Scientific Reports*, 8, 2018. doi: 10.1038/s41598-018-20214-6.

Wang, Z. Predicting or forecasting earthquakes and the resulting ground motion hazards: A dilemma for earth scientists. *Seismological Research Letters*, 86(1), 2015. doi: 10.1785/0220140211.

Ward, S. Area-based tests of long-term seismic hazard predictions. *Bulletin of the Seismological Society of America*, 85(5), 1995. doi: 10.1785/BSSA0850051285.

Wood, H. and Neumann, F. Modified Mercalli Intensity scale of 1931. *Bulletin of the Seismological Society of America*, 21(4), 1931. doi: 10.1785/BSSA0210040277.

Worden, C., Gerstenberger, M., Rhoades, D., and Wald, D. Probabilistic relationships between ground-motion parameters and modified Mercalli intensity in California. *Bulletin of the Seismological Society of America*, 102(1), 2012. doi: 10.1785/0120110156.